INF2178 Assignment 4
Mixed-Design ANOVA and Power Analysis on MMSE by Group and Visit

Liyanze Liu
1009172324

April 2024

1 Introduction

Cognitive evaluations play an essential role in the early detection and ongoing assessment of neurocognitive disorders. The Mini-Mental State Examination (MMSE) is a primary tool used in these evaluations, providing quantitative insights into cognitive function. In the analysis presented herein, MMSE scores are scrutinized using mixed-design ANOVA to discern patterns of cognitive performance across clinically distinct groups and across multiple assessment points. A subsequent power analysis confirms the statistical adequacy of the sample size, ensuring the results are both valid and reliable. The implications of this analysis are critical for understanding potential cognitive changes and for the design of future cognitive research.

2 Data and Methods

2.1 Data

This study uses a subset of longitudinal MRI data from the "INF2178_A4_data.csv" file, which captures the MRI results of patients with and without dementia. The data, chosen for its relevance to within-subject design and statistical power considerations, has been pre-selected for specific analyses.

2.2 Statistical Approach

2.2.1 Data Preparation and Visualization

The dataset was first cleansed of any records with missing 'SES' or 'MMSE' scores to ensure data quality. MMSE scores were then visualized to comprehend data distribution across clinical groups and visits.

2.2.2 ANOVA Framework

The primary analytical method employed was a mixed-design ANOVA, permitting the assessment of both between-subjects effects B related to the clinical group and within-subjects effects W corresponding to the visits. The hypotheses tested were:

- H_0: There is no interaction effect between the clinical group and visit on MMSE scores ($B \times W = 0$), and no main effect of either the clinical group ($B = 0$) or visit ($W = 0$).
- H_a: There is an interaction effect between the clinical group and visit on MMSE scores ($B \times W \neq 0$), or a main effect of either the clinical group ($B \neq 0$) or visit ($W \neq 0$).

ANOVA calculations were performed using the `pingouin` library, which facilitated a detailed assessment of these effects and their statistical significance.

2.2.3 Combined Statistical Analyses

Post-hoc t-tests discerned the specifics within group contrasts, revealing the subtle details of the main and interaction effects. The subsequent power analysis ensured a sample size adequate for detecting an effect size of 0.7, with a 0.91 power at an alpha of 0.05. Power curves illustrated the sample size against test power, aiding in understanding the likelihood of detecting the anticipated effect within dataset constraints.

3 Results

3.1 Visualization of MMSE Scores by Clinical Group and Visit

The point plot (Figure 1) of MMSE scores reveals discernible patterns across the two visits within each clinical group. For the nondemented group, scores remained relatively stable between the first and second visit. In contrast, the demented group showed a notable decline in MMSE scores over the two visits, suggesting a decrease in cognitive function. Interestingly, the converted group demonstrated an initial decline followed by a significant increase in MMSE scores, which may indicate variability in cognitive trajectories or responses to interventions.

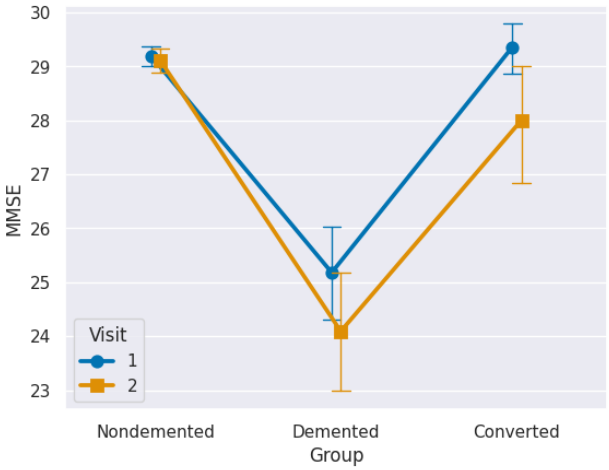

Figure 1: MMSE across clinical groups over visits

Group	Visit	MMSE	
		mean	std
Converted	1	29.36	0.93
	2	28.00	2.09
Demented	1	25.18	3.42
	2	24.09	4.49
Nondemented	1	29.19	0.85
	2	29.11	0.96

Table 1: Descriptive statistics of MMSE

3.2 Group Comparisons of MMSE Scores

Descriptive statistics computed from the MMSE scores indicate variability both within and between groups across two visits. For the Converted group, the mean MMSE score decreased from 29.36 (SD = 0.93) on the first visit to 28.00 (SD = 2.09) on the second, suggesting a decline in cognitive function. The Demented group exhibited a more pronounced decrease from 25.18 (SD = 3.42) to 24.09 (SD = 4.49), aligning with expectations of progressive cognitive impairment. Conversely, the Nondemented group's scores remained stable, with a mean of 29.19 (SD = 0.85) on the first visit and 29.11 (SD = 0.96) on the second, indicating consistent cognitive performance across time points (Table 1).

The observed standard deviations suggest greater variability in scores for the Demented group on the second visit, compared to the first. In contrast, the Converted group showed a notable increase in score variability from the first to the second visit, while the Nondemented group maintained relatively consistent variability across visits.

3.3 ANOVA

The two-way mixed-design ANOVA was conducted to evaluate the impact of clinical group status and visit occasion on MMSE scores (Table 2). The analysis revealed a statistically significant main effect for the clinical group, $F(2, 134) = 56.575, p < .001, \eta_p^2 = .458$, suggesting substantial differences in MMSE scores between the groups. The main effect for visit occasion was also significant, $F(1, 134) = 9.001, p = .003, \eta_p^2 = .063$, indicating changes in MMSE scores over time.

Furthermore, the interaction effect between clinical group and visit was found to be statistically significant, $F(2, 134) = 3.685, p = .028, \eta_p^2 = .052$. This suggests that the difference in MMSE scores across visits varied depending on the clinical group.

The results indicate that both the status of the clinical group and the number of visits play a crucial role in the cognitive performance measured by MMSE. Specifically, the significant interaction effect calls for a closer examination of how clinical group status may influence changes in MMSE scores over time.

Source	SS	DF1	DF2	MS	F	p-unc	np2	eps
Group	1325.778	2	134	662.889	56.575	0.000	0.458	nan
Visit	21.639	1	134	21.639	9.001	0.003	0.063	1.000
Interaction	17.716	2	134	8.858	3.685	0.028	0.052	nan

Table 2: ANOVA Summary

3.4 Post Hoc Comparisons of MMSE Scores

Following the ANOVA, post hoc tests were conducted to investigate the differences in MMSE scores between groups at each visit and the interaction effects between group and visit. The analysis revealed several significant contrasts:

- The difference in MMSE scores between the 'Converted' and 'Demented' groups was significant, $t(53.218) = 6.630, p < .001$, with a large effect size (Hedges' g = 1.185).
- When comparing the 'Demented' group across the two visits, there was a significant reduction in MMSE scores from visit 1 to visit 2, $t(57.340) = -9.032, p < .001$, indicating a decline over time with a large effect size (Hedges' g = -1.810).
- Notably, the interaction between visit and group was significant for the 'Converted' group from visit 1 to visit 2, $t(36.789) = -4.573, p < .001$, suggesting a differential trajectory of MMSE scores in this group relative to others (Hedges' g = 0.923).

These results suggest a distinct pattern of cognitive change for the 'Converted' and 'Demented' groups, with the 'Converted' group showing an increase in variability over time and the 'Demented' group exhibiting a consistent decline. The 'Nondemented' group did not show significant changes across visits, indicating stable cognitive performance over the study period.

3.5 Power Analysis for Sample Size Determination

A post-hoc power analysis was conducted to determine the adequacy of the sample size for detecting a true effect, if it exists. With a pre-defined alpha level of $\alpha = 0.05$ and a target power of $1 - \beta = 0.91$, the analysis sought to identify the required sample size for an anticipated effect size of $d = 0.7$. Utilizing the `TTestIndPower` tool from the statistical power analysis library, the results indicated that a sample size of approximately $n = 45.451$ participants per group would be necessary to achieve the desired power for detecting the specified effect size in a two-sided test.

This suggests that, for the study's results to have a high probability of detecting a true difference in MMSE scores between the groups if one exists, each group should consist of at least 46 participants (rounding up to the nearest whole number for practical implementation).

3.6 Power Curve Analysis

The power curve analysis was undertaken to determine how changes in sample size could influence the study's statistical power, given an effect size of 0.7 (Figure 2). The curve illustrates the relationship between the number of observations and the power of the test. Results indicated that with the effect size set at 0.7, the power of the test increased with the number of observations.

Specifically, the analysis revealed that a sample size of 20 yielded a power just under 0.80, which is commonly considered the minimum acceptable level. However, to achieve a more robust power of 0.90, a sample size of approximately 40 observations was required. The curve continued to ascend, approaching a power of 1.00 as the sample size approached 70 observations. This graphical representation confirms that the sample size calculated from the post-hoc analysis (approximately 45) would provide adequate power for this level of effect size.

Figure 2: Power of Test

4 Conclusion

This study's findings highlight significant differences in cognitive decline patterns among dementia, converted, and nondemented groups. A two-way mixed ANOVA indicated an interaction between group status and visit time on MMSE scores. The Converted group's scores dropped significantly, suggesting rapid cognitive decline, while the Nondemented group remained stable. The sufficient statistical power ensures these findings are reliable. Future studies should consider a larger sample size for generalizability. Understanding these patterns is vital for timely interventions in dementia care.